

Grundlagen

Was ist Visualisierung?

Visualisierung (auch Veranschaulichung) ist ein Tool um dem User Einsicht in Daten zu gewähren. Es werden abstrakte Daten (z.B Texte und Zahlen) und Zusammenhänge in eine graphische bzw. visuell erfassbare Form gebracht.

Visualisierungen reichen dabei von einfachen Grafiken bis hin zu komplexen medizinischen Veranschaulichungen.

Es handelt sich um Computergrafik, aber nicht um photorealistisches Rendering.

Beispiele:

- Geographische Karten (Kartographie), Wetterkarten.
- Zusammenhang zwischen Cholera Verbreitung und Wasserverfügbarkeit darstellen
- Röntgen
- Windsimulationen
- Simulationen

Folien:

- Tool to enable a User insight into Data
- To form a mental vision, image, or picture of (something not visible or present to the sight, or of an abstraction); to make visible to the mind or imagination.
- Computer Graphics, but not photorealistic rendering.

Womit beschäftigt sich VolVis, FlowVis und InfoVis?

VolVis:

veranschaulicht Volumendaten, hauptsächlich relevant für Medizinische Daten (Röntgen, Ultraschall). Ansonsten auch CAD Modellierungen für Industrie und Architektur.

FlowVis:

veranschaulicht "fließende" Daten, etwa Wasserströmungen, Luftbewegungen, Strahlen eines Düsenjets,...

InfoVis:

veranschaulicht abstrakte Daten. Etwa Trends, Statistiken, Zusammenhänge von ermittelten Werten,...

Welche 3 Ziele hat Visualisierung?

Erkunden/Erforschen (explore) von Daten:

Man weiß nichts über die Daten und muss diese erst erforschen.

Analysieren von Daten:

Zur Verifizierung oder Falsifizierung einer Hypothese.

Präsentation von Daten:

Die Informationen in den Daten sind bekannt. Visualisierung wird genutzt um das Wissen/die Ergebnisse zu kommunizieren.

Erklären Sie die Visualisierungspipeline.

1. Datengewinnung

- Messungen (zB CT/MRI)
- Simulation (zB Flow-Simulation)
- Modellierung (zB Game Theory)

2. Datenaufbereitung

- Filterung (zB smoothing/noise suppression)
- Resampling (zB andere Auflösung)
- Ableitung (zB Farbverläufe, Krümmung)
- Interpolation (zB linear, cubic)

3. Visualisierungsmapping (Daten in eine reederbare Struktur bringen)

- ISO-Oberflächenberechnung
- Glyphen, Icons festlegen
- Graphen Layout berechnen
- Voxel attributes: Farbe, Transparenz,...

4. Rendering

- Sichtbarkeitsberechnung
- Beleuchtung
- Composition (Kombinierung von transparenten Objekten)
- Animation

Wie hängt Datengewinnung mit der Darstellung zusammen?

Je nach Art der Daten kommen unterschiedliche Visualisierungsmodelle zum Einsatz. Bei der Visualisierung fokussiert sich alles auf die Veranschaulichung der Daten.

Was sind Daten? Warum sind sie wichtig?

Daten beinhalten (so hofft man) Information, welche durch Visualisierung sichtbar wird. Ohne Daten kann nichts visualisiert werden. Unterschiedliche Typen von Daten erfordern unterschiedliche Visualisierungstechniken.

Wo „leben“ die Daten?

Räumliche Daten:

In 2D und 3D Räumen (SciVis).

zB bei medizinischen Daten, Strömungssimulationsdaten, GIS-Daten,...

Nicht räumliche Daten:

Abstrakte Daten. Räumliche Einbettung durch Visualisierung. zB Datenbanken, Statistiken.

Aspekte: Dimensionalität (data space), Koordinaten, Einflussregion (local, global), domain.

Welche Typen haben Daten?

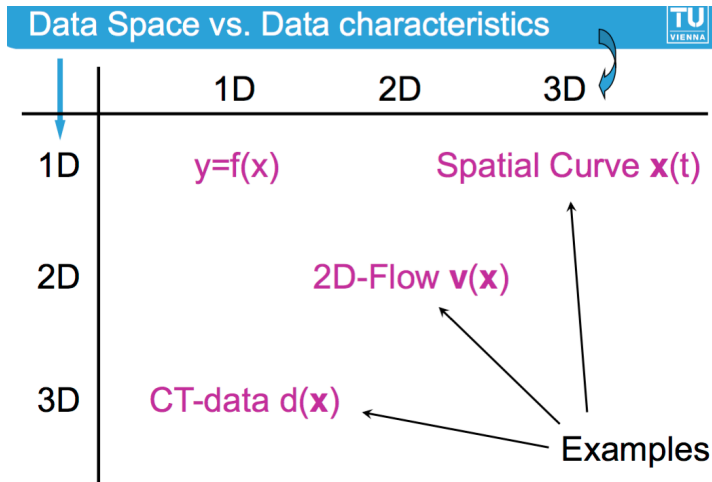
- Scalar: numerische Werte (natürliche, ganze, rationale, reale, komplexe Zahlen)
- nicht Numerisch (nominale, Ordinate Werte)
- Multidimensionale Werte (n-dimensionale Vektoren, n*n-dimensionale Tensoren)
- Multimodale Werte (Vektoren mit Daten von variierendem Typ, zB Zeilen in einer Tabelle)

Wie kann man Daten darstellen?

Wenn sich die Daten in einem Raum befinden kann dieser wiederverwendet werden (2D, 3D-Darstellungen). Dies muss aber nicht sein (zB Histogramm eines Bildes).

Wahl der Dimension:

- Beziehung zwischen Datenraum und Daten-Eigenschaften



- Verfügbarer Darstellungsraum (2D/3D)
- Wo liegt der Fokus?
- Wo kann abstrahiert werden / gespart werden (zB zu viele Dimensionen)

Wie kann man Mengendaten darstellen?

Eulerdiagramm, Venndiagramm, Radial Sets, Parallel Sets.

Was ist ein Grid und wozu wird es verwendet?

ein Gitter beschreibt...

- ...die Messpunkte (Anzahl mit der ein Objekt abgetastet wird).
- ...wie Daten organisiert sind.
- ...die Nachbarschaftsbeziehung zwischen den Punkten.
- ...wie Navigation innerhalb des Gitters möglich ist.
- ...die Struktur der Daten.
- ...ob Berechnungen möglich sind.

Welche Arten von Grids kennen Sie?

Kartesisches Gitter

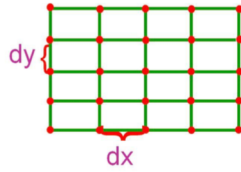
- Orthogonales Gitter mit einheitlichen Entfernungen.
- Uniforme Distanzen (in allen Dimensionen gilt: $dx = dy$)
- Implizite Nachbarschaftsbeziehung.

Regular Grid – Rectilinear Grid



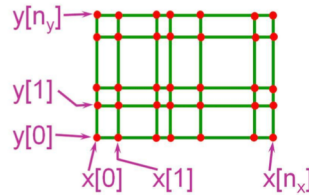
Regular Grid

- ◆ $dx \neq dy$



Rectilinear Grid

- ◆ varying sample-distances $x[i], y[j]$

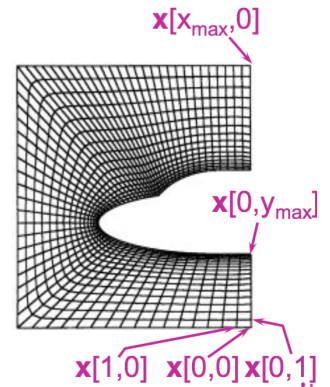


Curvilinear Grid



Characteristics:

- ◆ non-orthogonal grid
- ◆ grid-points explicitly given $(\mathbf{x}[i,j])$
- ◆ Implicit neighborhood-relationship

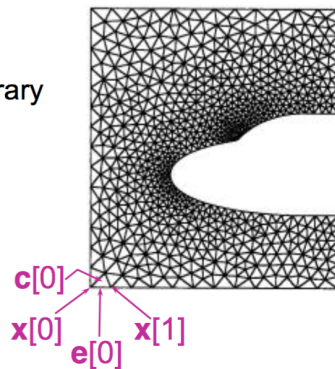


Unstructured Grid

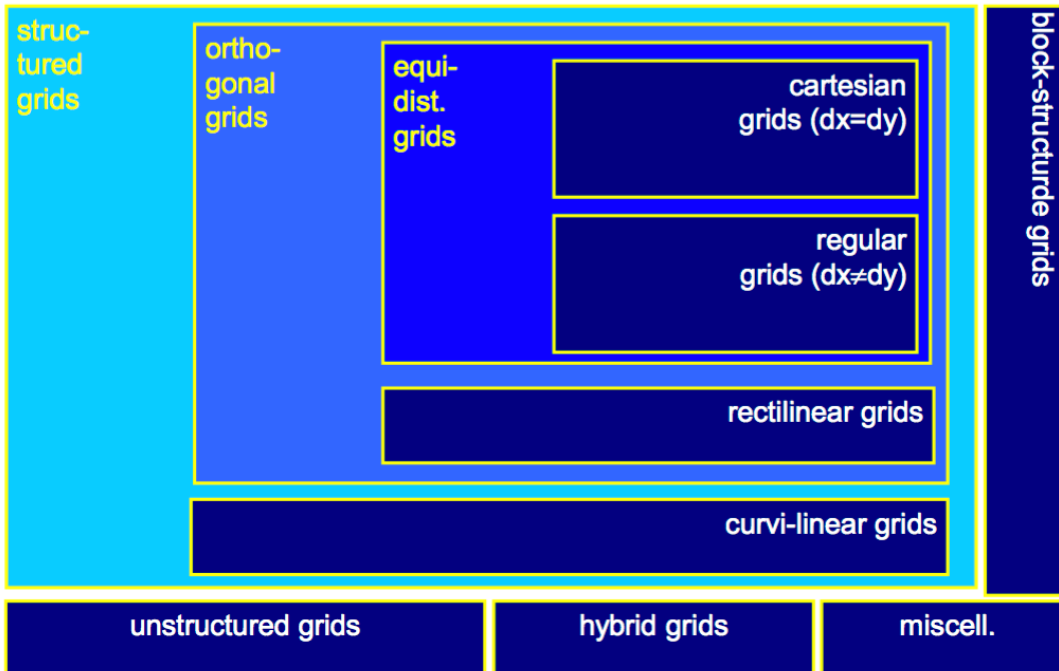


Characteristics:

- ◆ Grid-points and connections arbitrary
- ◆ Grid-points and neighborhood explicitly given
- ◆ Cells: tetrahedra, hexahedra



Grids - Survey



Was sind scattered data?

Eigenschaften von Scattered Data:

- haben kein Gitter
- haben keine Nachbarschaftsbeziehung
- Einfluss auf die Nachbarschaft durch räumliche Nähe definiert.
- Scattered data interpolation

Grid Transformations

Es kann vorkommen, dass man zwischen Gittertypen wechseln muss.

Die Umwandlung kann stattfinden auf:

- der Physikalischen Domain (Simulation)
- der computational Domain (visualization Mapping)
- Bild Domain(Rendering)

Wichtige Faktoren für die Grid Transformation sind:

- die Genauigkeit des Resampling
- und das Design von Algorithmen.

Welche Regeln gelten bei der Verwendung von Farben bei einer Visualisierung?

- Verwenden von entsättigten Linien als Rand von farbigen Flächen.
- Kein sattes Blau für Details und Animationen verwenden.
- Nie gesättigtes Blau und Rot mischen
- Vermeiden von hohen Farbfrequenzen
- Farben zum vergleichen sollen nah zueinander sein
- Beachten von Kontext und Assoziationen.
- Verwendung von Redundanz (Form, Stil, etc.)

Fakten über Farbe:

- Farbe kann Informationen hervorheben
- Es sollten maximal 7 +- 2 Farben verwendet werden um eine bessere Überschaubarkeit zu gewährleisten.
- Es können ca. 50-300 Schattierungen einer Farbe erkannt werden (abhängig von der Farbe)
- Die Regenbogen-Farbskala ist nicht linear
- Die Farbwahrnehmung ist abhängig vom Kontext
- Farbblinde Benutzer sind gehandicapt (auf Barrierefreiheit achten wenn möglich)

Volumenvisualisierung

Was ist Volumenvisualisierung?

Unter Volumenvisualisierung versteht man das Darstellen von 3D Daten in einer 2D Projektion.

Woher stammen die Volumendaten und wie sind sie strukturiert?

Volumendaten stammen von:

- Medizinischen Anwendungen (Computer Tomographie, Magnetresonanz)
- Materialtests (Industrie-CT)
- Simulationen(Finite Element Methods, Computational fluid dynamics)
- Modellen (Architektur, Auto-, Flugzeugs-, Schiffsmodelle)

Zur Strukturierung wird gerne das Cartesian oder Regular Grid herangezogen. Diese eignen sich gut für die Speicherung von CT und MR Daten.
Ansonsten werden auch Cervi-Linear Grids verwendet, welche als tetrahedra oder hexahedra organisiert sind.

Welche Probleme gibt es bei Volumenvisualisierung?

- bei der Render-Projektion kann es zu Problemen kommen, da so viel Information vorhanden ist, aber nur begrenzt viele Pixel zur Verfügung stehen.
- Es entstehen sehr schnell sehr große Datenmengen. Etwa wäre ein $512 \times 512 \times 1024$ Voxel-Datensatz zu je 16bit schon 512MB groß.
- Die Geschwindigkeit der Darstellung sollte über 10FPS liegen, um eine Echtzeit-Interaktion zu ermöglichen.

Was ist der Unterschied zwischen Voxel und Cell?

Ein Voxel ist ein Volumenelement, welches einen Samplewert beschreibt. Diese Werte werden nicht notwendigerweise interpoliert.

Eine Cell ist ein Würfelprimitiv, das durch seine 8 Ecken beschrieben wird. Jede Ecke wird von einem Voxel beschrieben, die Werte innerhalb der Cell werden interpoliert.

Welche Interpolationen kennen Sie? Beschreiben Sie diese.

Nearest Neighbour: Der Wert aus dem nächstgelegenen Voxel wird für eine Position herangezogen.

Trilinear Interpolation: Es wird ein Wert aus allen umliegenden Voxel interpoliert, welche mit ihrer Entfernung zur Position gewichtet werden.

Was ist ein Gradient?

Gradienten ersetzen die Normalvektoren, da keine Flächen sondern nur Punkte vorhanden sind. Gradienten zeigen in die Richtung des höchsten Anstiegs

Wie kann man die Gradienten berechnen?

Der Gradient wird berechnet indem die erste Ableitung der Funktion gebildet wird und ein Punkt eingesetzt wird.

Was ist Klassifikation und wozu wird sie verwendet?

Bei der Klassifikation geht es darum verschiedene Bereiche im Datensatz zu klassifizieren. Etwa soll bei einem Datensatz eines menschlichen Körpers zwischen Haut, Muskeln und Knochen unterschieden werden können.

Dabei wird zB auf Dichtewerte, Gradienten und Krümmung zurückgegriffen.

Um eine automatische Klassifikation vorzunehmen werden Transferfunktionen verwendet.

Was ist eine Transferfunktion?

Transferfunktionen bieten eine vorgefertigte Klassifikation. So werden unterschiedlichen Dichtewerten unterschiedliche Farben zugeordnet. (zB Luft=Schwarz, Haut=Gelb, Knochen=Rot)
Allgemein: Daten werden Farb- und Transparenzwerte zugeordnet.

Was ist der Unterschied zwischen Slice, Surface und Volume Rendering?

Slice Rendering: es wird nur eine Scheibe des Datensatzes gerendert. So bekommt man Einsicht in ein Objekt, als ob es durchgeschnitten wurde.

Surface Rendering: Die Oberfläche eines Gegenstands wird mittels ISO-Surface gerendert.

Volume Rendering: Die Renderfunktion ist abhängig von der Transferfunktion.

Welche Volumenvisualisierungstechniken kennen Sie?

Slicing, ISO-Surface, Ray Casting.

Erklären Sie Slicing.

Beim Slicing werden axenparallele Scheiben dargestellt. Durch diese kann man im Regelfall durchscrolen um unterschiedliche Stellen zu sehen.

Slicing benötigt eine Transferfunktion um den Dichtewerten des Datensatzes Farbwerte zuzuordnen.

Was ist der Unterschied zwischen Image-Order und Object-Order?

Bei Ray Casting wird bei Image-Order für jedes Pixel ein Strahl auf das Objekt geschickt. Der Berechnungsaufwand hierfür ist von der Auflösung des Bildes abhängig.

Bei Object-Order werden vom Objekt die Strahlen auf die Bildebene geschickt. Das Objekt wird hierfür umgedreht. Der Berechnungsaufwand ist von der Objektgröße abhängig.

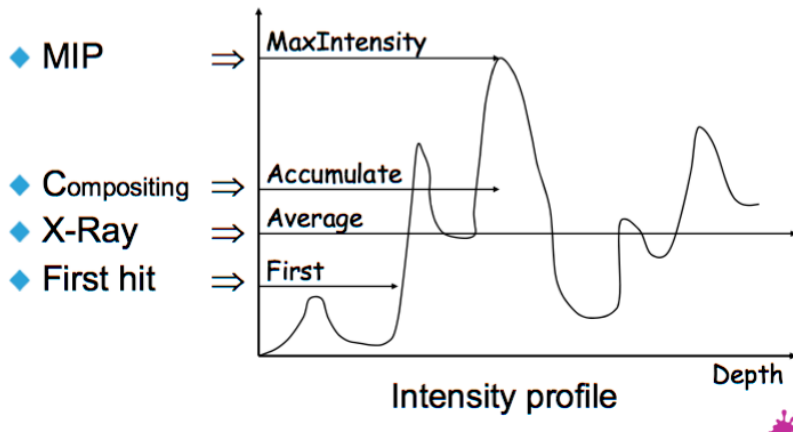
Erklären Sie Ray Casting.

Beim Ray Casting werden ausgehend von der Bildfläche Strahlen auf das Objekt geschickt und dadurch die notwendigen Dichte- oder Farbwerte ermittelt.

In der Praxis wird beim Casten des Strahls durch das Objekt Ray Traversal angewendet. Das heißt es werden in gewissen Abständen diskrete Samples genommen. (dies kann in genormten Abständen stattfinden, ein mal per durchschrittenen Voxel oder per Voxelübergang ein mal.

Es gibt dabei verschiedene Kombinationen wie die Farb/Dichtewerte zum tragen kommen:

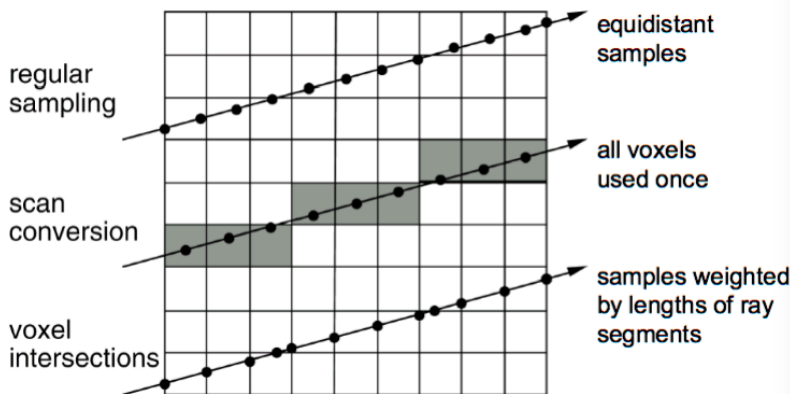
Overview:



Was ist Ray Traversal?

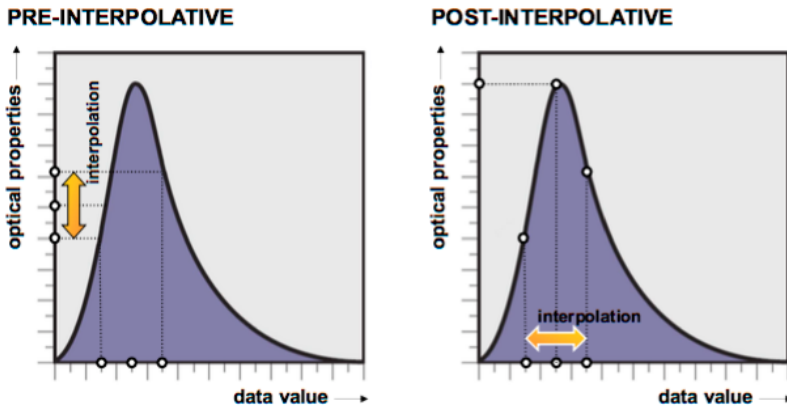
Ray Traversal ist das Durchschreiten des Volumendatensatzes. Es werden dabei in gewissen Schrittabständen diskrete Samples genommen.

2. Ray Traversal – Three Approaches



Was ist der Unterschied zwischen pre- und post-interpolative Klassifikation?

bei der pre-inerpolation wird zuerst gesamplet und dann interpoliert.
 bei der post-interpolation wird zuerst interpoliert und danach klassifiziert.



Erklären Sie den physikalischen Hintergrund von α -compositing.

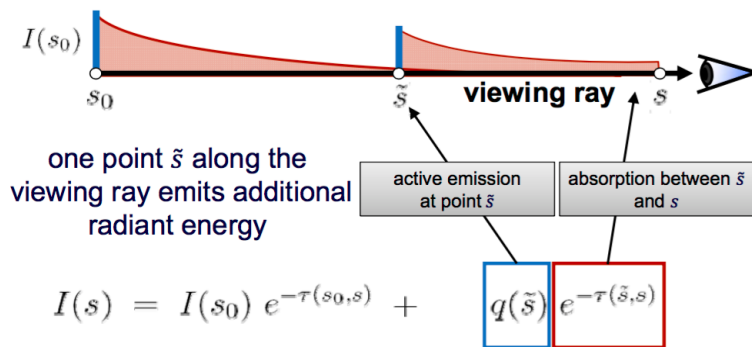
Mit Alpha-Compositing werden Transparenzeigenschaften von Materialien nachgestellt. So ist es möglich bei halb transparenten Materialien die Gegenstände dahinter durchzusehen. Es können auch Emission und Absorption von Licht nachgestellt werden.

- Emission zB bei Licht Partikeln
- Absorption zB bei dunklen Nebel
- Emission und Absorption bei Wolken

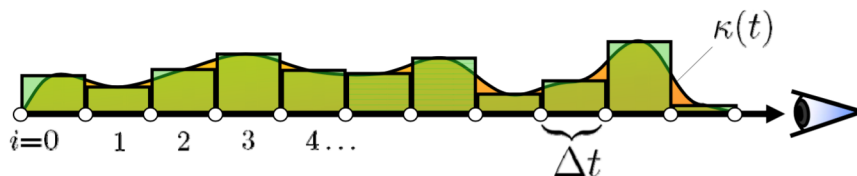
Erklären Sie α -compositing.

Analytisches Modell:

- Licht schwächt sich über die Distanz ab $e^{-\tau}$
- Je höher die Anzahl der Partikel, desto höher ist die Abschwächung * Hinterer Teil wird zusätzlich addiert(Integral)



Numerische Approximation:



- Dichtewert entlang des Strahls wird mittels Riemannsummen approximiert. —> Integral wird zur Summe
- Messpunkte werden gleichmäßig verteilt
- Es wird Farbe und Durchsichtigkeit approximiert

- Back to Front Compositing: Die Voxel werden vom Hintergrund zum Betrachter durchlaufen. Der für die vorherigen Voxel berechnete Farbwert muss jeweils mit der Transparenz und der Farbe des aktuellen Voxels verrechnet werden (akkumuliert werden). Funktioniert wie ein Paintersalgorithmus → man sieht den Aufbau.
- Front to Back Compositing: Die Voxel werden vom Betrachter (Augpunkt) aus zum Hintergrund hin durchlaufen. Die Farbe weiter hinten liegender Voxel geht entsprechend kaum noch in die Gesamtfarbe ein, so dass der Algorithmus bei Unterschreitung eines bestimmten Schwellenwertes abgebrochen werden kann. Man muss sich merken welche Durchsichtigkeit und welche Farbwerte es gibt, es muss sich also mehr gemerkt werden. Ergebnisse sind bei beiden Verfahren gleich, beide können als Alphablending bezeichnet werden.
- Numerische Lösungen: Die Farbe kann mit dem Alphawert vor multipliziert werden. → Vereinfachung (Problem: Transparentes Rot == Transparentes Schwarz) → Verlust von Präzession

Was bedeutet pre-multiplied Alpha?

Die Farbwerte (RGB) werden bereits vorab mit dem Alphawert multipliziert (aR, aG, aB). Dies hat zur Folge, dass transparentes Rot das gleiche ist wie transparentes Blau oder transparentes Schwarz.

Mit dem vorab Multiplizieren wird dafür das Blending erleichtert, da Farbe und Alphawerte gleich behandelt werden. Gleichzeitig kann aber auch ein Verlust an Genauigkeit auftreten.

Welche 2 Aspekte kennen Sie bei Hardware Visualisierung?

3D texturen:

Die Volumendaten sind in der Textur gespeichert. Beim Rendering werden die Volumendaten dem Viewport angepasst. Es findet ein Back-to-front Compositing statt.

2D Texturen:

Bestehen aus 3 Stacks aus Scheiben (x-, y- und z-Achse). Es wird zum Rendern immer der Stack ausgewählt, welcher am parallelsten zur Bildfläche ist. Es findet ebenfalls Back-to-front Compositing statt.

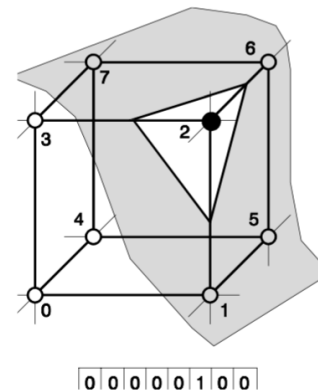
Was sind ISO-Flächen?

ISO-Flächen sind Flächen, die im Raum benachbarte Punkte gleicher Merkmale oder Werte einer bestimmten Größe wie zum Beispiel Temperatur oder Dichte miteinander verbinden.

Erklären Sie den Marching Cubes Algorithmus?

Marching Cubes

- * Gestalt der Zelle nur aufgrund der Eckpunkte
- * ist ein Algorithmus zur Darstellung von Isoflächen in der 3D-Computergrafik. Er nähert eine Isofläche durch eine Polygongrafik an.
- * Die Idee von Marching Cubes ist es, das gegebene Voxelmodell eines Objekts zunächst in kleine Würfel (cubes) zu zerlegen und anschließend von einem Würfel zum nächsten zu „marschieren“ und zu bestimmen, wie die Oberfläche des Objekts den jeweiligen Würfel durchschneidet.
- * Man vergleicht die Eckpunkte mit den Isowerten, darüber oder darunter.
- * Algorithmus:
 1. Erstelle einen Würfel
 2. Klassifiziere jeden Voxel, überprüfe ob Fläche innerhalb des Würfels ist: Wert > Iso Wert -> Punkt liegt ausserhalb
 3. Erstelle einen Index - verwende binäres Labeling innerhalb = 1, außerhalb = 0, es gibt 256 verschiedene Möglichkeiten wie Würfel durchlaufen werden kann. -> das kann auf 15 vermindert werden.(durch Ausnützung von Symmetrie)
 4. Lookup Edge List 256 auf 15 reduzieren, man kann mit 0 bis 4 Dreiecken auskommen.
 5. Oberflächenschnittpunkt für jede Dreiecks-Kante interpolieren. Man bekommt innerhalb der Zelle den Ort an dem der Schnittpunkt liegt.
 6. Berechne die Normale für jede Ecke(Vertex) des Würfels
 1. Problemfälle: zweideutige Fälle: es können Löcher entstehen, Lösung: Entscheidung für einen Fall. Man bekommt konsistente Lösung(keine Löcher) aber es können Fehler gemacht werden. Oder man löst es exakt -> Asymptotische Entscheidung(Hyperbeln werden zur Entscheidung verwendet, Signum), Es gibt insgesamt 6 Problemfälle.



Welche Nachteile hat der Marching Cubes Algorithmus?

Es können Löcher in Objekten entstehen. Dies kann passieren da mehrere mögliche Würfel für einen Fall existieren. Das Problem kann gelöst werden indem ein konsistenter Fall gewählt wird, was aber zu Fehlern im Rendering führen kann.

Strömungsvisualisierung

Was ist Strömungsvisualisierung?

Strömungsvisualisierung macht verschiedene Strömungen sichtbar. Etwa Wasserströmungen, Luftströmungen von Düsenjets, Wetter,...

Strömungsvisualisierungen werden sowohl in 2D als auch 3D realisiert. Es gibt gleichmäßige Strömungen und zeitabhängige Strömungen. Strömungen können indirekt und direkt visualisiert werden.

Am einfachsten ist eine Visualisierung mit Pfeilen, ansonsten können auch Strömungslinien oder Partikel verwendet werden.

Woher stammen die Strömungsdaten?

Simulationen

- Flugzeug-, Schiff-, Autodesign
- Wettersimulation (Luft, Wasser)
- Medizin (Blut, Körperflüssigkeiten)

Messungen

- Windtunnel, Flüssigkeitstunnel
- Schlieren- und Schattentechniken

Modelle

- Differentialgleichungssysteme(ODE)

Was sind die Eigenschaften der 2D, 3D und Flächenvisualisierungen?

- 2D Visualisierungen: Zeigen eine Scheibe aus einer 3D Realität. ZB Wolken- und Luftströmungen.
- 3D Visualisierungen: Zeigen Bewegungen in drei Dimensionen, zB für Simulationen in der Industrie.
- Flächenvisualisierungen: Zeigen Bewegungen auf der Oberfläche eines Gegenstandes. Es handelt sich also um einen Gegenstand, der von Strömungen umgeben ist. Dabei wird gezeigt an welchen Stellen die Strömungen ihn wie betreffen. zB Oberfläche eines Flugzeugs.

Was ist der Unterschied zwischen steady und time-dependent Strömungen?

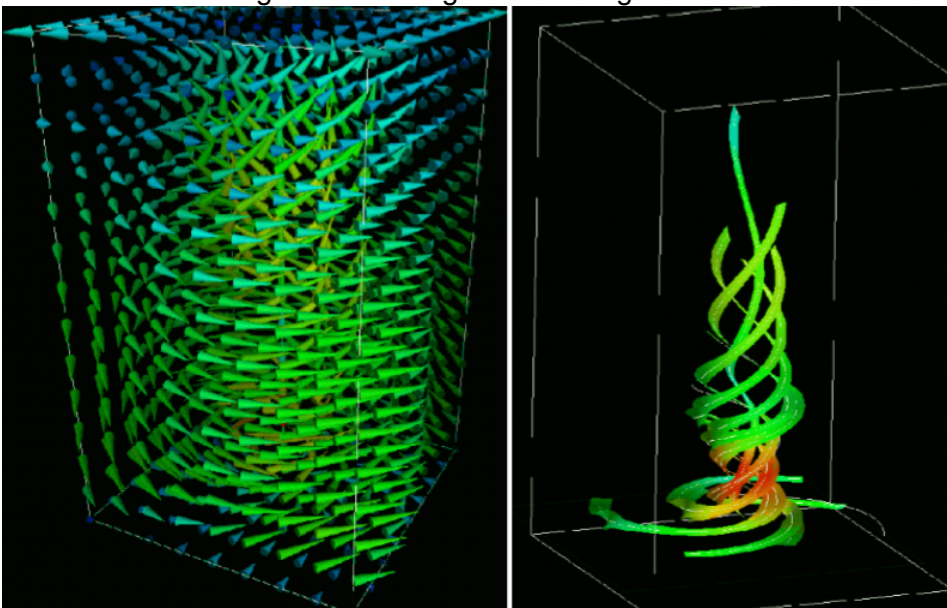
Steady Flow: Die Strömung ändert sich nicht. Einfacher zu visualisieren. (Gitter mit Richtungsvektoren)

Time-dependent Flow: Die Strömung verändert sich im Lauf der Zeit. Komplexere Visualisierungsmodelle werden benötigt. (Hubschrauber-Rotor, Wetterprognose)

Was ist der Unterschied zwischen direkter und indirekter Strömungsvisualisierung?

Direkt: Die Strömung wird direkt dargestellt. Das heißt an jeder Stelle (bzw. in regelmäßigen Abständen) wird die Strömung visualisiert (zB mit Pfeilen oder Smearingtechnique)-

Indirekt: Es werden Objekte wie Strömungslinien berechnet und dargestellt. So kann auch eine zeitliche Entwicklung der Strömung statisch dargestellt werden.



Welche experimentelle Strömungsvisualisierungsmethoden kennen Sie?

- Injektion von Farbe, Rauch oder Partikeln in Strömungen.
- Optische Methoden: Schlieren, Schatten

Was ist PIV?

Particle Image Velocimetry.

Die Geschwindigkeit der Partikel werden zu zwei aufeinanderfolgenden Zeitpunkten gemessen. Dazu werden die Partikel durch ein Sieb möglichst gut gestreut. Dann werden mit einer Highspeedkamera zwei Aufnahmen nacheinander gemacht, welche mit einem Stroboskop beleuchtet werden.

Wie kann man Modelle Visualisieren?

Einfache Lösung: Handskizzen

Visualisierung mit Pfeilen.

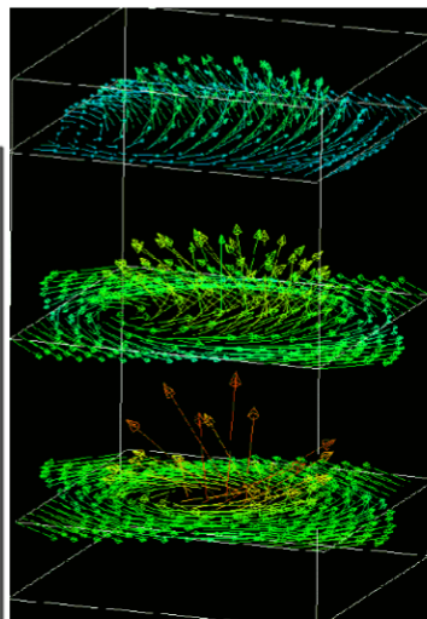
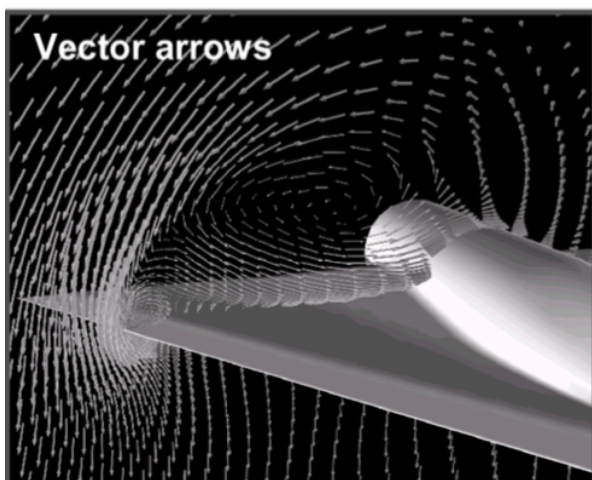
Eigenschaften:

- Direkte Flussvisualisierungen
- Normalisierte Pfeile oder Skalierung mit Flussgeschwindigkeit.
- 2D oft einfacher zu verwenden als 3D, was Probleme bereiten kann.
- Oft limitierte Aussagekraft, da die zeitliche Komponente nicht dargestellt werden kann.
- gerne verwendet, da einfach.

3D Pfeile:

- Mehrdeutigkeit,
- Perspektivische Verkürzung,
- Verbesserbar durch 3D Pfeile
- Für bessere Wahrnehmung kann die Strömung auf eine "Scheibe" reduziert werden.

■ **Compromise:** **Arrows only in slices**



Ansonsten mit Strömungslinien.

Erklären Sie die Ziele der Visualisierungen mit Pfeilen.

- Einfache Darstellungsweise.
- Direkte Art Strömungen zu visualisieren.
- Pfeile können entweder normalisiert oder mit der Geschwindigkeit skaliert dargestellt werden.
- Besonders gut für 2D Darstellungen geeignet.
- Im 3D Bereich können Probleme durch Mehrdeutigkeit und perspektivische Verkürzungen auftreten.
- Farbcodierung möglich.

Was sind Strömungslinien und wozu werden sie verwendet?

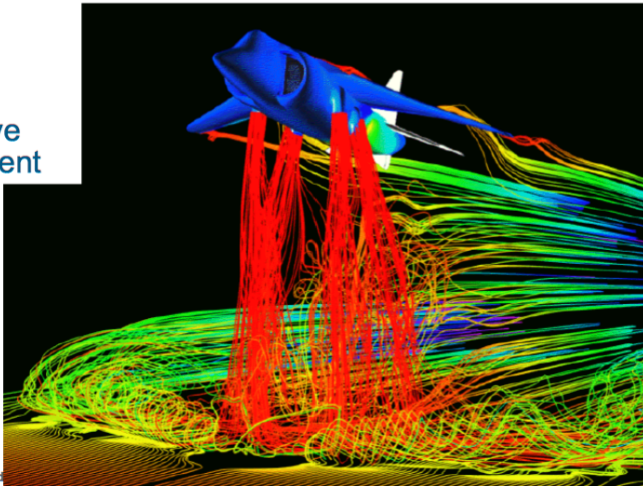
Die Strömung kann anhand von Linien visualisiert werden -> Strömungslinien.

Diese sind als Kurve definiert, wobei an jeder Stelle die Tangentialrichtung dem zugrundeliegenden Strömungsfeld entspricht.

Strömungslinien bieten im Gegensatz zu Pfeilen eine bessere Übersicht, wie sich die Partikel weiterbewegen.

Um die Bewegungsrichtung zu visualisieren, kann die Spitze der Linie dicker gezeichnet werden.

- Color coding: Speed
- Selective Placement



Helwig Hauser, Eduard

Wie berechnet man die Strömungslinien?

Es wird ein Seed-Point gewählt. Von diesem ausgehend wird die Tangentialrichtung des zugrundeliegenden Strömungsfeldes ausgerechnet, an welcher das Partikel weiterwandert. Dies wird in regelmäßigen (kleinen) Abständen wiederholt um eine Linie zu generieren.

Erklären Sie die Euler'sche Integration.

Man macht eine Linearisierung und nimmt an dass sich in einem ganz kleinen Zeitraum der Partikel nicht bewegt. —> damit verkommt das Integral zu einer Summe(Euler Integral), Je kleiner die Schrittweite gewählt wird desto kleiner ist der Fehler(bsp. Ellipse). Problem: Fehler wird immer in eine Richtung gemacht —> Extrembeispiel ist die Ellipse, bei Kurven sind die Ergebnisse besser.

Bei einer Schrittweite dt von $1/100$ ist der letzte Punkt wieder fast im Ursprung.(2% entfernt), man benötigt aber 889 Schritte.

Erklären Sie das Runge-Kutta-Verfahren.

Die Verbesserung von Euler ist der Runge Kutta Ansatz. Der Fehler wirkt sich im Quadrat aus (Runge-Kutta Verfahren 2. Ordnung). Dafür ist das Verfahren aufwändiger, man verwendet einen sogenannten Preview Vektor. Der Vektor wird in der Mitte unterteilt.

1. Man macht einen halben Eulerschritt
2. Evaluieren den Flussvektor.
3. Verwendet ihn im Ursprung

Runge Kutta ist mit 9 Schritten ($dt = 1$) besser als Euler mit 72 Schritten ($dt = 1/9$)

Nennen Sie Beispiele für Strömungsvisualisierungen mit Strömungslinien.

- Wetterkarten (Hurrikan)
- Stromlinie um Auto/Flugzeug
- Stromlinien von Senkrechtstarter

Wie platziert man die Strömungslinien?

Würde man die Seed-Points gleichmäßig auf einem Gitter verteilen, würde dies zu einer sehr unregelmäßigen Verteilung der Strömungslinien führen.

Deswegen verwendet man einen Füllalgorithmus, der die Linien gleichmäßig verteilt. Dazu legt man eine Distanz fest. Von einem beliebigen Seed-Point ausgehend zeichnet man eine Strömungslinie mit dem Abbruchkriterium falls eine andere Strömungslinie näher als die festgelegte Distanz ist.

Abbruchkriterien:

- wenn die Linie einer anderen Linie näher als die festgelegte Distanz kommt.
- wenn außerhalb des Darstellungsraumes,
- wenn Strömung in Fixpunkt endet (Strömung = 0),
- wenn sich die Strömungslinie selber zu nahe kommt,
- wenn eine maximale Anzahl an Schritten erreicht wurde.

Nennen Sie Beispiele für Strömungsvisualisierungen mit Integral-Objekten.

Es werden zum Veranschaulichen der Strömungen nicht einfache Partikel verwendet, sondern Linien oder sogar Flächen. So ist es möglich ein Streamribbon (Strömungsband) zu visualisieren, welches auch Verwirbelungen darstellen kann.

Streaklines werden zB bei Autos verwendet, um zu sehen wie sich die Partikel verhalten können. Flow Volumes werden verwendet um Luftströmungen zu zeigen, oder wie sich zwei Massen ineinander vermischen,

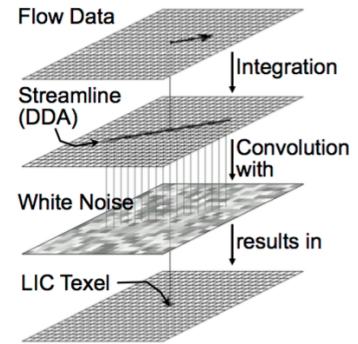
Was ist Line Integral Convolution?

Bei LIC geht es darum Strömungen als Textur zu visualisieren. Die Erstellung funktioniert ähnlich wie bei Tintentröpfchen die auf weißes Papier aufgetragen werden und dann vom Wind verweht werden. Die Farbe (oder bei LIC das weiße Rauschen) wird entlang der Linien des

Strömungskanals verteilt und mit einem Weichzeichner (Gauß) gefiltert.

Line Integral Convolution - Strömungsvisualisierung in 2D oder auf Oberflächen

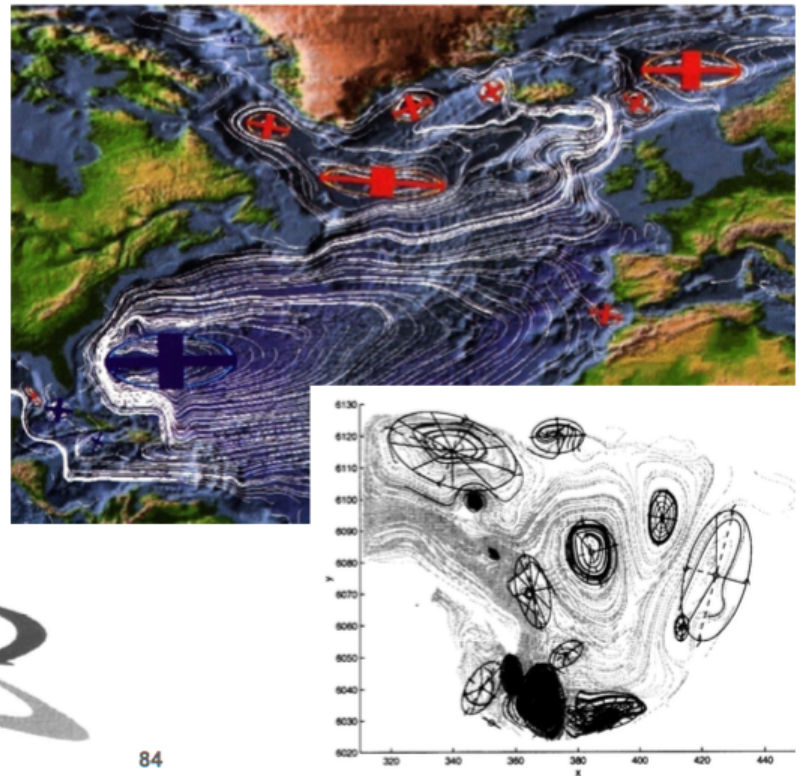
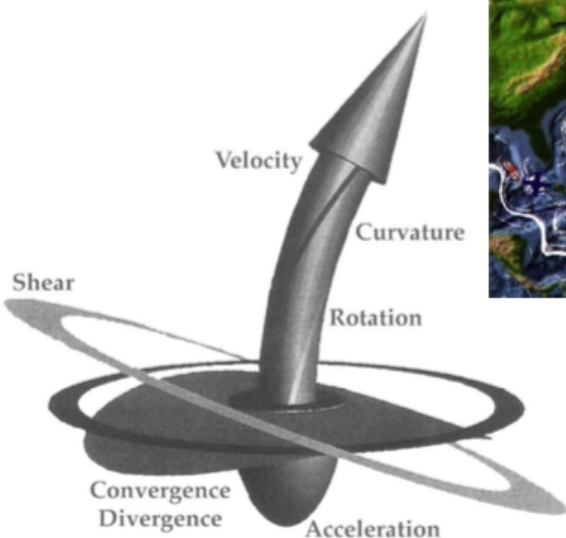
- * In 3D nicht wirklich gut anwendbar, wegen Überdeckung
- * Strömung wird als Textur visualisiert
- * Feldrichtungen werden überall angezeigt
- * Funktioniert wie Tintentröpfchen(White Noise) die auf Blatt Papier aufgetragen werden. Oder Eisenspäne mit Magnet
- * Um eine LIC-Darstellung für solch ein Strömungsfeld zu erhalten, wird ein zweidimensionales weißes Rauschen entlang der Linien des Strömungsfeldes mit einem Kern mit lokalem Träger gefaltet (im einfachsten Fall: lokal gemittelt). Außerdem wird das Bild mit einem Weichzeichner(Gauß Filter) gefaltet.
- * Die Pixelwerte entlang der Feldlinien sind dadurch stark korreliert, orthogonal dagegen - durch das Rauschen - fast unkorreliert. Dadurch heben sich die Feldlinien optisch vom Hintergrund ab und werden sichtbar.
- * Kann auch in 3D gemacht werden, ein Problem ist aber die Verzerrung
- * Alternativen: Spot-Noise(Ellipsen werden entlang der Strömung gelegt), Vector Kernel, Line Bundles/Splats(man hat Textursplats)
- * Nachteil: Aufwändig, Richtung wird nicht dargestellt.



Nennen Sie Beispiele für Strömungsvisualisierungen abhängig von lokalen Eigenschaften.

Bei einer Strömung können verschiedenste Parameter vorhanden sein. zB Druck, Rotationsverhalten, Beschleunigungsverhalten, Verwirbelungen. Durch verschiedene Visualisierungstechniken möchte man diese veranschaulichen. zB durch Ringe und Scheiben, Länge, Orientierung oder Farbe.

Local / topological properties



84

Was ist die Topologie einer Strömung?

- Eine abstrakte Struktur der Strömung.
- Es zeigt die wichtigsten Eigenschaften der Strömung aber nicht Details.
- Es können dabei Zyklen entstehen.
- Checkpoints und Verbindende Strukturen können verwendet werden. (auf Oberflächen Fixpunkte und Verbindungen)
- Lorenz System: ist eine 3D Darstellung von Flow-Topologie.
- Timesurface: Veranschaulicht eine Menge von Punkten wo nicht die räumliche, sondern die zeitliche Entwicklung im Vordergrund steht.

Informationsvisualisierung

Was ist Informationsvisualisierung?

“The use of computer-supported, interactive, visual representations of abstract data to amplify cognition”

Das bedeutet wir verwenden Computer um Information so darzustellen, dass Menschen sie leichter verstehen können und einen besseren Überblick darüber erhalten.

Wozu wird Informationsvisualisierung verwendet?

Exploratory Data Analysis:

Man hat keine Hypothese zu den Daten. Die Daten werden durchsucht und analysiert um darin potenziell nützliche Informationen zu finden. Mit diesen Informationen wird eine Hypothese aufgestellt.

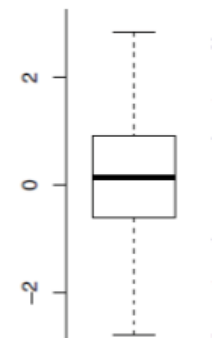
Diese Methode wurde von John Tukey(1915-2000) eingeführt, der ebenfalls den Box-Plot erfunden hat.

Confirmatory Analysis:

Als Ausgangspunkt dient eine Hypothese, diese wird mit Hilfe von Visualisierung auf dessen Korrektheit überprüft. Als Ergebnis erhält man ob die Hypothese korrekt oder falsch ist.

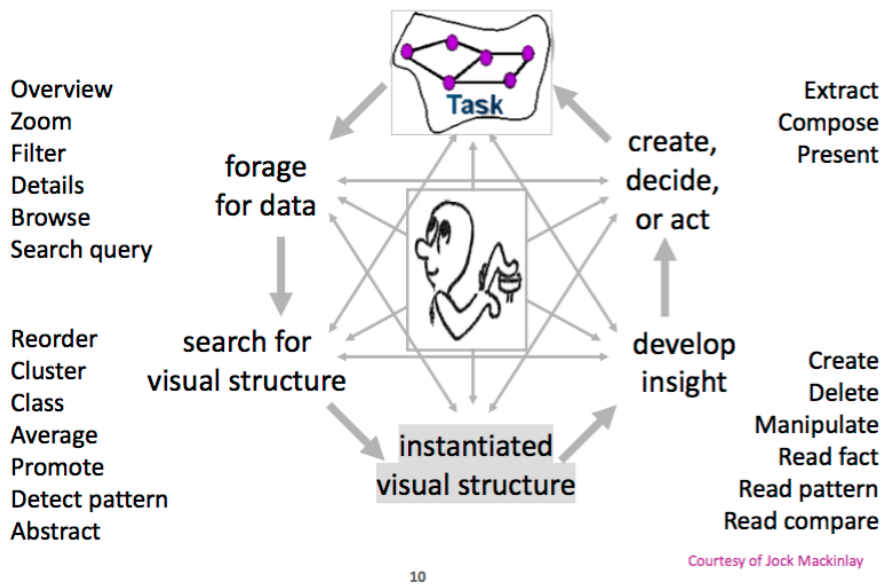
Presentation:

Die Fakten existieren bereits. Hier ist das Ziel dieses Wissen anderen so effizient und effektiv wie möglich zu präsentieren.



Erklären Sie den Knowledge-Crystallisation-Loop.

Es existiert eine Aufgabenstellung, für diese wird nach den passenden Daten gesucht. Die Daten werden in eine passende Struktur gebracht, mit welcher man versucht einen möglichst guten Überblick über die Informationen zu gewinnen. Mit der erhaltenen Einsicht werden weitere Maßnahmen gesetzt wie etwas erstellen, Entscheidungen treffen oder handeln (wieder am Anfang der Loop)



10

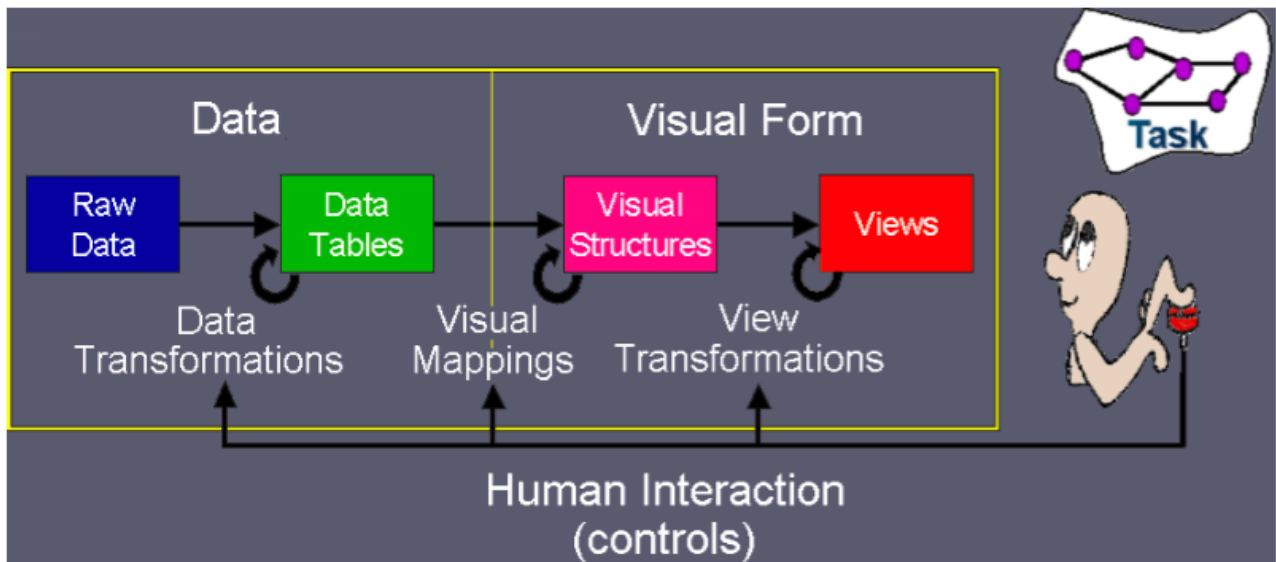
Courtesy of Jock Mackinlay



Ein Beispiel hierfür ist der Dynamic HomeFinder, es werden zur Verfügung stehende Wohnungen angezeigt, welche nach Kriterien gefiltert werden können (Gebiet, Kosten, Größe, Zimmeranzahl). Nach dem Suchen können die Ergebnisse sortiert werden (Kosten, Größe, Zimmer,...)

Erklären Sie das InfoVis Referenzmodell.

Es wird von den vorhandenen/gewonnenen Daten ausgegangen. Diese werden mittels Datentransformationen in Tabellen geordnet und dann mit Visual Mappings in Visuelle Strukturen gebracht (Bäume, Grids, Netzwerke). Die visuellen Strukturen werden weiter in Views gebracht, welche ausgegeben werden und dabei gefiltert, sortiert oder verzerrt werden können. Auch Details on Demand möglich.



- Raw Data: idiosyncratic formats
- Data Tables: relations(cases by variables)+metadata
- Visual Structures: spatial substrates + marks + graphical properties
- Views: graphical parameters (position, scaling, clipping, zooming,...)



Raw Data Probleme:

- Fehler, Inkonsistente Darstellung
- Variablenformate
- Fehlende Daten
- Variablentypen inkonsistent
- Unterschiedliche Tabellenstrukturen passen nicht

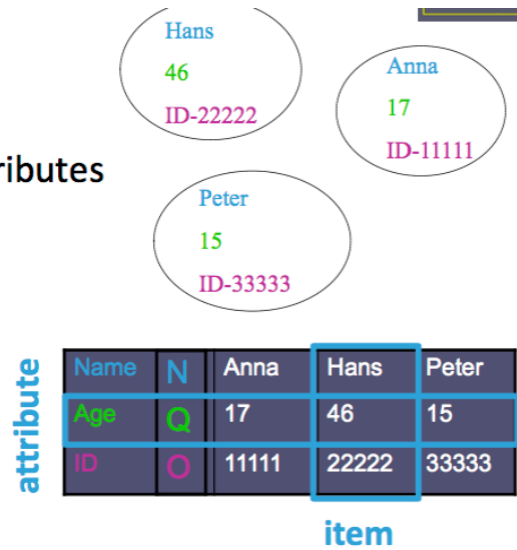
Data Transformation:

- Der Prozess um Rohdatenmaterial in Tabellen zu bringen. Wird verwendet um Datentabellen zu verbessern oder Produzieren (zB aus Texten).

Datentabellen:

■ **Data Tables:**

- ◆ Cases / Items
- ◆ Variables / Attributes
 - Nominal
 - Quantitative
 - Ordinal
- ◆ Values
- ◆ Metadata



Daten Transformation:

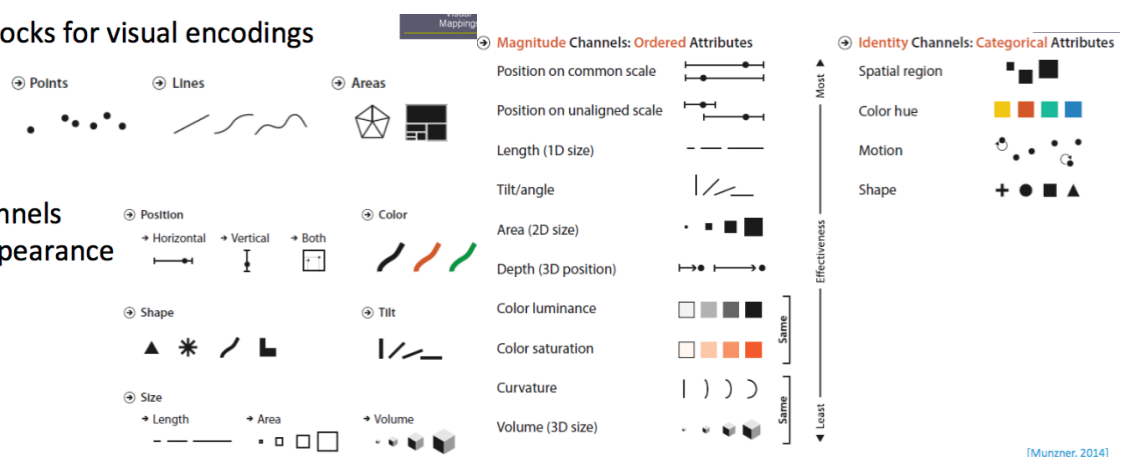
- Aus Werten werden Werte und/oder Strukturen abgeleitet.
- Aus Strukturen werden Werte und/oder Strukturen abgeleitet.

Visual Mappings/Visual Structures:

■ **Building blocks for visual encodings**

■ **Marks:**

■ **Visual channels control appearance of marks**



[Munzner, 2014]

Hier muss immer beachtet werden was gezeigt werden soll. zB: Soll der Anteil vom Ganzen gezeigt werden, eignet sich ein Kreisdiagramm. Für einen Vergleich eignet sich ein Stabdiagramm besser.

Für Multivariate Daten bieten sich Scatterplots an (oder Bubble Charts zB gapminder), eine Scatterplot Matrix, Parallele Koordinatensysteme oder Radar Charts.

View Transformation/View:

Wenn statistische Daten in Visualisierungen umgewandelt werden.

Ansätze sind:

- Overview and Detail (zB Große Detailkarte + kleine Übersichtskarte)
- Zooming (Je weiter in die Karte gezoomt wird, desto mehr Details werden eingeblendet)
- Focus and Context (zB Kartesische Fisheye-Transformation: das Gitter wird an der Stelle größer wo sich der Mauszeiger befindet. So kann man die Details an der Stelle besser betrachten)

(Folien Beachten, sehr viele Darstellungsmöglichkeiten!)

Welche Eigenschaften haben Trees?

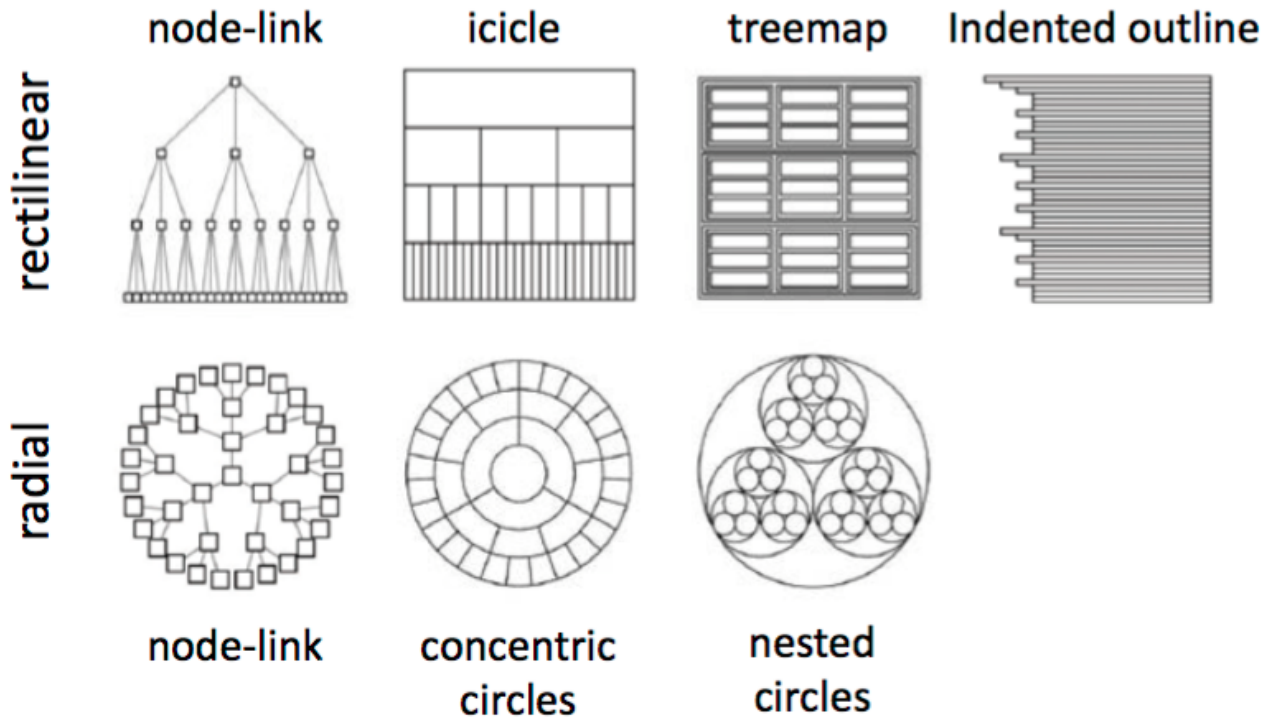
Bäume sind visuelle Strukturen, welche Verbindungen Zwischen Knoten und Blättern darstellen, um Beziehungen zwischen Daten darzustellen.

- Bäume repräsentieren hierarchische Strukturen.

Wünschenswerte Eigenschaften:

- Planarität (keine Überschneidung von Kanten)
- Klarheit (Die Beziehungen über Kanten sollten deutlich erkennbar sein = möglichst kurze und gerade Kanten).
- Klares, nicht verschlungenes Design.
- Hierarchisches Design sollte in eine Richtung zeigen.

Welche Arten von Trees kennen Sie?



[Munzner, 2014]

Welche Eigenschaften haben Netzwerke?

Netzwerke sind gut geeignet um Kommunikationsnetzwerke wie Telefonsysteme oder das Internet oder U-Bahnsysteme zu veranschaulichen.

- Beispiel U-Bahn in Knotenpunkten wird mehr Platz beansprucht, deswegen werden die Daten verzerrt. Außerdem sind die Richtungen verändert. Wichtig ist aber das die Topologie vorhanden bleibt.
- Metabolische Netzwerke: Netzwerke die zeigen wie Stoffe vom Körper verarbeitet werden. (sehr komplex, man kann zB Kanten weglassen)

Knoten sind:

- Unstrukturiert
- Nominal
- Ordinal
- Quantitativ

Verbindungen sind:

- Gerichtet oder ungerichtet

Netzwerke können als Node-Link Diagramm oder Adjazenzmatrix dargestellt werden. Für größere Netzwerke kann man diese auch kombinieren (Matrizen werden zu Knoten in Node-Link Diagrammen)

Was ist eine visuelle Transferfunktion?

Eine Transferfunktion, die eine Visualisierung so verzerrt (vergrößern oder verkleinern an gewissen Stellen), dass relevante Regionen genauer dargestellt werden, um mehr Details zeigen zu können. DOI - Degree Of Interest Function

Wann ist Interaktion wichtig?

- Um dynamische Abfragen abzusetzen
- Details on Demand
- Brushing (Es werden die selben Daten in unterschiedlichen Diagrammen dargestellt. Es kann ein Datensatz ausgewählt werden, welcher in allen Diagrammen hervorgehoben wird)

Kann zB umgesetzt sein durch Checkboxen, welche Inhalte relevant sind oder anklicken auf Punkte um Details zu erhalten.

Visual Analytics

Was ist Visual Analytics?

Visual Analytics ist die Wissenschaft analytisches Denken mit einem Visuellen Interface zu unterstützen. Es werden hierfür die automatische Analyse mit Techniken der interaktiven Visualisation kombiniert, um ein effektives Verstehen, Begründen und Entscheidungen-Treffen auf Basis von sehr großen und komplexen Datenmengen zu ermöglichen. "Detect the expected and discover the unexpected"

zB eingesetzt bei

- Geheimdiensten um Terroristen zu finden
- Unwetter wie Hurricanes um zu evaluieren wo Barrieren errichtet werden sollen.

- Physik und Astronomie: Das Erkennen von unerwarteten Phänomenen in riesigen und dynamischen Datenströmen.
- Katastrophenschutz: Die Analyse einer Notsituation, um geeignete Gegenmaßnahmen zu entwickeln, die helfen den Schaden einzugrenzen (Naturkatastrophen, etc.).
- Biologie und Medizin: Die Analyse großer Mengen an Bio-Daten (das menschliche Genom, etc.).
- Business-Intelligence: Analyse von Kundendaten

Wikipedia:

Visual Analytics ist ein interdisziplinärer Ansatz, der die Vorteile aus unterschiedlichen Forschungsgebieten verbindet. Das Ziel der Visual-Analytics-Methode ist, Erkenntnisse aus extrem großen und komplexen Datensätzen zu gewinnen. Der Ansatz kombiniert die Stärken der automatischen Datenanalyse mit den Fähigkeiten des Menschen, schnell Muster oder Trends visuell zu erfassen. Durch geeignete Interaktionsmechanismen können Daten visuell exploriert und Erkenntnisse gewonnen werden. Er wurde 2004 eingeführt und ein Jahr später in dem Buch "Illuminating the Path" beschrieben.

Erklären sie den Visual Analytics Prozess.

1. Daten werden berechnet und transformiert. (Gesäubert, normalisiert, gruppiert, fusioniert)
Data: Heterogene Datenquellen müssen vor der visuellen oder automatischen Analyse zuerst vorverarbeitet werden (z.B. bereinigt, normalisiert, etc.).
2. Models: Mit Hilfe von Data-Mining-Techniken werden Modelle der Originaldaten generiert, welche daraufhin zu Evaluierungszwecken oder für weitere Verbesserungen visualisiert werden.
3. Visualization: Um die Modelle durch einen Benutzer zu überprüfen, werden Visualisierungen generiert, welche mit Interaktionstechniken für eine Analyse angereichert werden.

Die Vorgehensweise orientiert sich dabei an folgendem Paradigma:

„Analyse First – Show the Important – Zoom, Filter and Analyse Further – Details on Demand“
erweitert von: “Overview first, zoom/filter, details on demand”.

Dabei ist ein stetiger Wechsel zwischen visuellen und automatischen Vorgängen eine wichtige Eigenschaft des Visual-Analytics-Prozesses. Verfälschte Resultate können dadurch frühzeitig erkannt werden, um ein besseres und vertrauenswürdigeres Endergebnis zu erhalten.

Welche Herausforderungen hat Visual Analytics?

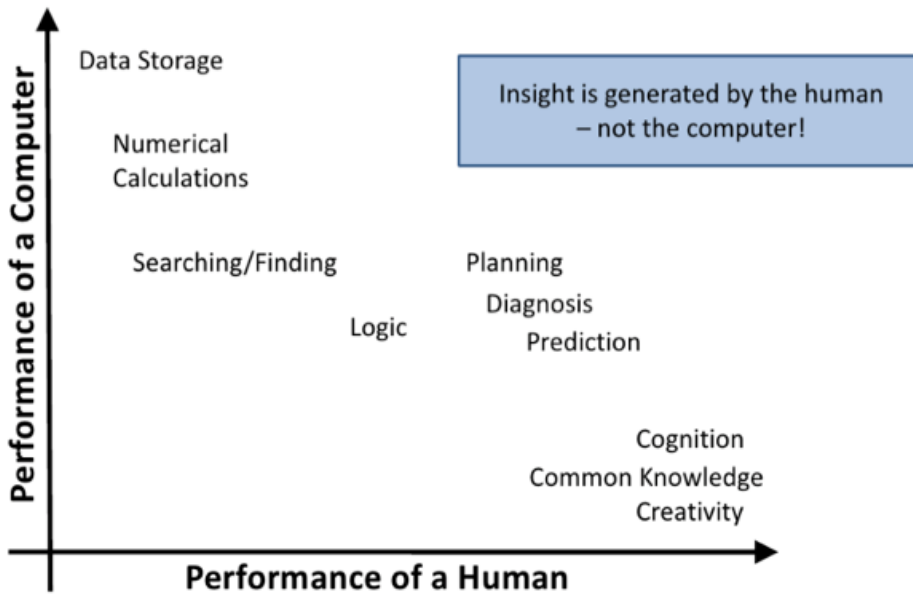
- Data: es wird mit extrem großen, diversen und unterschiedlich großen Datensets gearbeitet.
- Users: Die Bedürfnisse der Benutzer müssen befriedigt werden
- Design: Das Design soll Designern von Visual Analytic Systemen Helfen?!
- Technologie: Die benötigte Infrastruktur zur Verfügung stellen.

Was ist Data Mining?

Aus den Rohdaten werden mit Hilfe von Algorithmen nützliche Informationen extrahiert. Dies kann Semiautomatisiert oder Vollautomatisiert stattfinden.

Was ist die Ability Matrix?

Die Ability Matrix beschreibt welche Aufgaben ein Computer besser erfüllen kann und welche vom Menschen besser erledigt werden können.



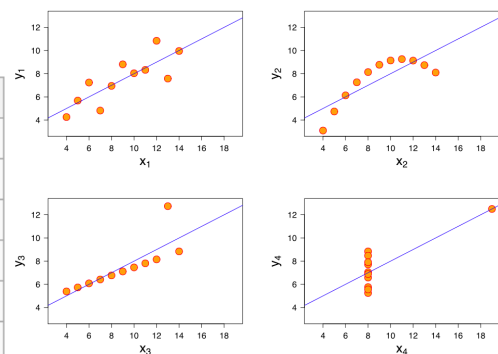
Warum verwendet man Graphiken um die Daten zu visualisieren?

Weil der Sehsinn der am besten ausgeprägte Wahrnehmungssinn des Menschen ist. Mit Abbildungen lassen sich mehr Informationen übersichtlicher auf weniger Platz darstellen. Abbildungen bieten den Gestalteffekt, welcher eine schnelle Übersicht bietet und mehr Struktur in Informationen bringt. Figuren sind leichter zugänglich, leichter zu verstehen, schneller zu begreifen, einprägsamer, bereiten mehr Spaß und sind weniger formell.

Was ist das Anscombe-Quartett?

Das Anscombe-Quartett besteht aus vier Mengen von Datenpunkten, die nahezu identische einfache statistische Eigenschaften haben, aber aufgetragen sehr verschieden aussehen. Jede dieser vier Mengen besteht aus elf (x,y)-Punkten. Diese vier Mengen wurden im Jahre 1973 von dem englischen Statistiker Francis Anscombe konstruiert, um die Bedeutung einer graphischen Datenanalyse herauszustellen und die Effekte von Ausreißern zu demonstrieren

Eigenschaft	Wert
Mittelwert von x in jedem Fall	9 (exakt)
Varianz von x in jedem Fall	11 (exakt)
Mittelwert von y in jedem Fall	7,50 (auf 2 Stellen)
Varianz von y in jedem Fall	4,122 oder 4,127 (auf 3 Stellen)
Korrelation zwischen x und y in jedem Fall	0,816 (auf 3 Stellen)
Lineare Regression in jedem Fall	$y = 3,00 + 0,500x$ (auf 2 bzw. 3 Stellen)

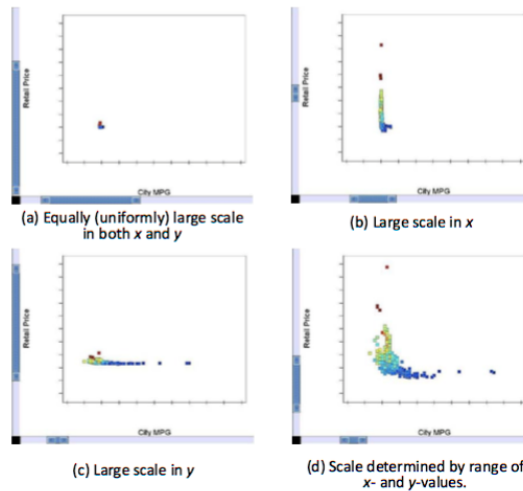


Welche Probleme hat man bei verzerrten Visualisierungen?

Die Darstellung kann einem Bias unterliegen. Je nach gewählter Skala können die Daten komplett unterschiedliche Eigenschaften suggerieren.

Visualization Can Be Biased

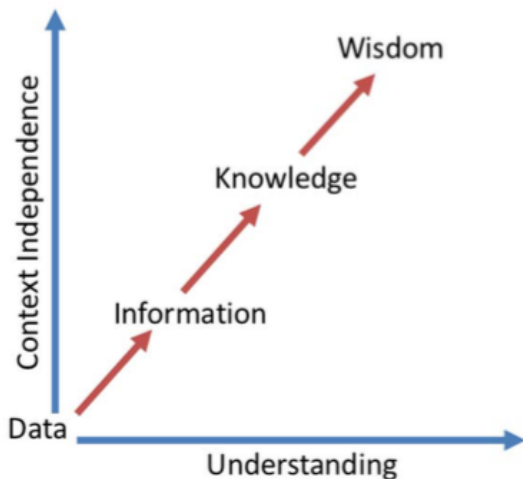
The same data plotted with different scales is perceived dramatically differently.



Was ist der Unterschied zwischen Diagramm und Visualisierung?

Ein Diagramm repräsentiert Information.
Eine Visualisierung repräsentiert Daten.

Was ist die DIKW-Hierarchie? Erklären Sie die Begriffe Data, Information, Wissen und Weisheit.



The DIKW-Hierarchy according to [Bellinger 2004]

From Epistemology:

Data = ground truth

Information = phenomena

Knowledge = causes

Wisdom = possible (inter-)actions

"Above all else, show the data"

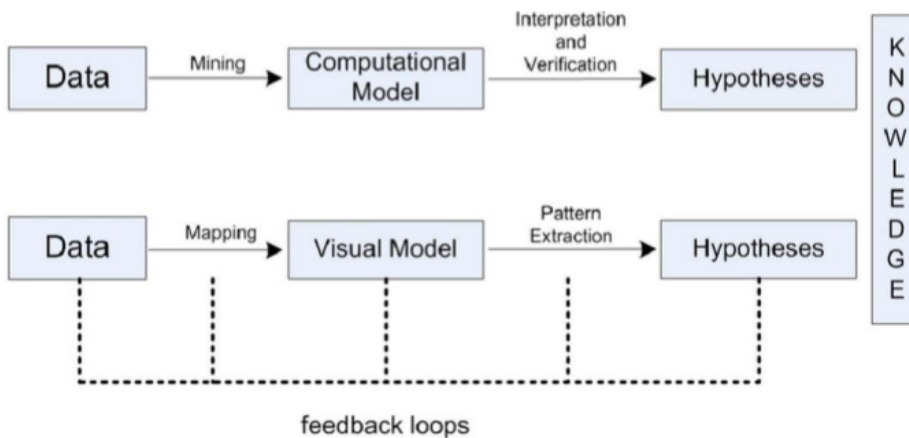
- Edward R. Tufte

Was ist das Visual Analytics Mantra?

Analyse first, show the important, zoom/filter, analyse further, details on demand.
erweitert von: "Overview first, zoom/filter, details on demand".

Was ist der Unterschied zwischen Data Mining und Visual Analytics Prozessen?

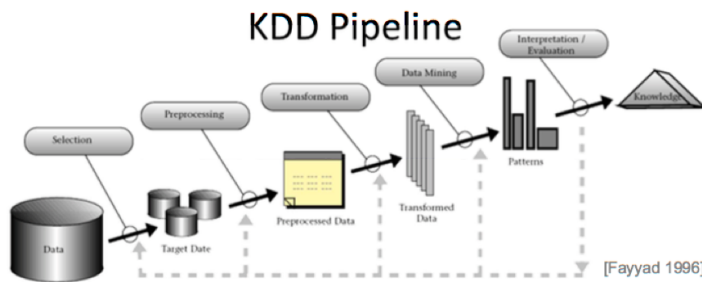
Bei Datamining werden aus den Daten (semi-)automatisch Informationen gewonnen. Bei Visual Analytics werden die Daten visualisiert und aus der Visualisierung Wissen gewonnen.



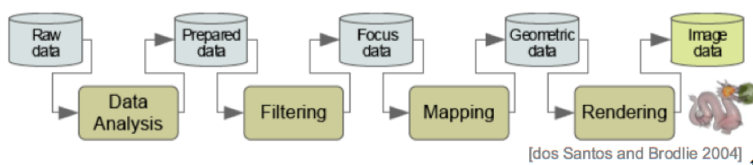
Was ist der Unterschied zwischen der KDD und der Visualisierungspipeline?

Knowledge Discovery in Databases

Der Unterschied besteht hauptsächlich darin, dass bei KDD auch die Bewertung der Resultate inkludiert ist.



Visualization Pipeline



Was ist Unsicherheit bei Visual Analytics?

- What is not surrounded by uncertainty cannot be the truth [Richard Feynman]
- True genius resides in the capacity for evaluation of uncertain, hazardous, and conflicting information [Winston Churchill]
- Doubt is not a pleasant condition, but certainty is absurd [Voltaire]

-> Es muss immer Unsicherheit berücksichtigt werden, diese ist immer zu einem gewissen Grad vorhanden. Es können allerdings trotzdem Rückschlüsse gezogen werden. In manchen Bereichen möchte man Unsicherheit ausblenden, zB Arzt bei seinen Patienten.

Ungenauigkeit durch Messgeräte, Unmessbarkeit, können schon in Daten eliminiert werden, oder erst später.

Welche Herausforderungen hat Data Management?

Was ist „Big Data“?